Monkey detection using deep learning for monkey-repellent

Nur Latif Azyze Mohd Shaari Azyze¹, Teow Khimi Quan¹, Ida Syafiza Md Isa², Muhammad Afif Husman³

¹Department of Mechatronics Engineering, Fakulti Teknologi Dan Kejuruteraan Elektrik, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

²Department of Electronic and Computer, Fakulti Teknologi Dan Kejuruteraan Elektronik Dan Komputer, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

³Department of Mechatronics, Kulliyyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia

Article Info

ABSTRACT

Article history:

Received Jun 7, 2024 Revised Nov 28, 2024 Accepted Jan 14, 2025

Keywords:

Convolution neural network Deep learning Repellent Supervised learning YOLO Animal intrusion has caused many issues for human beings, especially monkeys. Monkeys have caused many problems such as yield crop damage, damage to human facilities and assets and stealing food. This study aims to investigate the use of deep learning to detect monkey presence accurately and use a proper repellent system to scare them away. A deep learning algorithm is constructed with supervised learning, which includes the monkey dataset with appropriate labelling of the object of interest. The detection of the monkey comes with a bounding box with respective confidence of detection. The result is then used to evaluate the accuracy of monkey detection. The accuracy of the trained model is assessed by evaluating its performance under varying conditions of camera quality and distances. The study focuses on proving the reliability of deep learning to detect monkeys and automatically perform corresponding actions like repelling monkeys. Hence this may reduce the reliance of manpower to secure a large space and improve safety issues.

This is an open access article under the <u>CC BY-SA</u> license.



Corresponding Author:

Nur Latif Azyze Mohd Shaari Azyze Fakulti Teknologi Dan Kejuruteraan Elektrik, Universiti Teknikal Malaysia Melaka Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia Email: latifazyze@utem.edu.my

1. INTRODUCTION

Human-monkey conflict has been a serious issue over the years for example damage to yield crops, intrusion into human facilities and stealing human food. The wildlife department has continued to cull long-tailed macaques, with up to 70,000 animals being killed annually between 2013 and 2016 due to the average of 3,800 complaints from the public made nationally each year. However, according to [1] the culling of these monkeys has decreased in recent years, with officials now focusing more on transferring rather than killing them.

Moving the monkey will not solve the problem in the long term. But it cannot be sure that other groups of monkeys will not come again in the future. This is where animal classification comes in particularly handy to solve human-monkey conflicts in the long term [2]. In these circumstances, an image classifier/identifier is useful since it automates the process of classifying and identifying the monkey [3]. Utilizing deep learning integrated with a convolutional neural network (CNN) approach facilitates the straightforward detection of monkeys via camera [4]. This method is also presented in other applications which resulted in the object being detected successfully [5]–[7]. This study requires a computer that can run Python code for training the model algorithm, then it can interact with a 24/7 camera for monitoring the presence of a monkey and a speaker to produce appropriate repellent sound to scare the monkey away [8], [9].

Animal intrusion has been a problem for human beings for a long time. Previous researchers has been taking part in safeguarding the assets and life of human beings which include the residential area, working place, institution, and agricultural crop area [10]–[13]. The most common animal related to this kind of intrusion is the monkey.

Conventional animal detection with passive infrared (PIR) sensors lacks accuracy and visual evidence, potentially compromising reliability [14]. These sensors detect presence but not location precisely, sensitive to factors like object size and distance. This can affect transparency and reliability, crucial for repelling monkeys while excluding other moving objects. Additionally, the machine learning model's effectiveness in identifying monkeys may be hindered by inadequate training data, leading to biased or poor performance [15]-[19]. The monkey detection method refers to the various techniques and approaches used to identify and locate monkeys in each environment [20], [21]. The line model approach consists of several processes to detect monkeys, the process is background extraction, star skeletonization, and line model matching process [22], [23]. A straightforward background subtraction technique is used to remove the primary constituent portion from the background. Supervised learning algorithms are trained with labelled examples, such as an input with a known desired output [24]. When presented with new data, the model is trained to recognize the underlying patterns and correlations between the input data and the output labels, allowing it to produce accurate labelling results. However, the accuracy of monkey detection has not yet been studied across video qualities and at varying distances. This is because the accuracy of the model is subject to the input image quality [25]. The main objective of this project is to develop an image classification method using YOLOv7 algorithm tailored for accurately detecting monkeys. The accuracy of this detection will be assessed across various levels of image quality and distances, which are critical factors influencing the effectiveness of the model in practical scenarios.

2. METHOD

An integrated pipeline for monkey detection using deep learning is illustrated in Figure 1. First, the collection of the dataset is obtained by capturing an image at different areas with different angles. Next, each image is labelled to create a comprehensive monkey cluster for training purposes. The data was trained by the YOLOv7 algorithm to educate the machine learning on recognizing monkeys. Following the training phase, the accuracy and performance of the trained model are evaluated. If the results do not meet expectations, hyperparameter tuning is conducted, adjusting parameters such as learning rate, optimizer weight decay, and momentum among others. In this study, 363 monkey images are used with 70, 20 and 10% for training, validation and test respectively. Follows are details explanations of dataset collection, dataset labelling and training model.

2.1. Dataset collection

In the context of monkey video analysis, the process involves recording a monkey video using a smartphone or camera and then converting it into a set of separate photos. The dataset collection and preparation process as follows: i) begin by recording a video of monkeys using a smartphone or camera, ensuring the video quality is sufficient to capture their behavior and movement effectively while considering factors such as lighting conditions, camera stability, and subject clarity; ii) after recording the video, proceed to convert it into a sequence of individual images using an MP4 to JPG converter available online. This tool extracts each frame from the video and saves it as a separate JPEG image file. It is important to select a dependable converter that supports your specific video format and offers essential customization features; iii) after converting the video into an image sequence, there will be a series of frames representing different moments from the recorded video. Not all frames may be suitable for the dataset, especially those lacking significant movement. It is important to manually review and select frames that show noticeable changes from one frame to another. This ensures that the dataset includes diverse instances of monkey behavior, facilitating effective learning by the model; and iv) evaluate the quality of the selected frames to confirm they meet the dataset criteria. Factors to consider include clarity, focus, lighting, and overall image quality. Exclude frames that are blurry, distorted, or inadequately lit to maintain a dataset comprising high-quality images. This process aims to reduce noise and unwanted variations that could hinder the model's training effectiveness.

2.2. Dataset labeling

Dataset labelling is done using the application "*labelImg*", *LabelImg* is a popular graphical image annotation tool used for labeling objects in images. It provides a user-friendly interface for manually annotating objects of interest. "*Open Dir*" is where to place our dataset which is ready for label, and "*Change Save Dir*" is the directory where we save the annotation file. "*Create RectBox*" is used to create a rectangular

bounding box, once we draw the bounding box, a pop-out window will prompt us to input the label for the image, in this case "monkey".

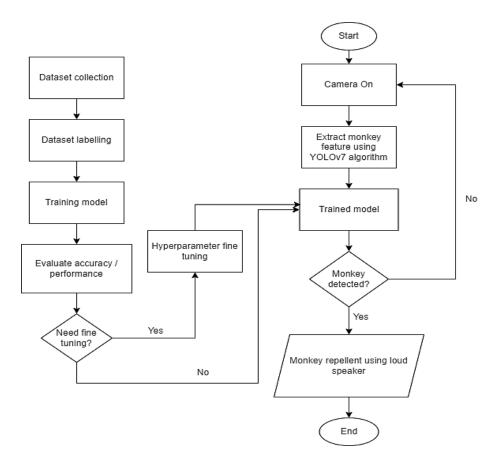


Figure 1. Integrated pipeline for monkey detection using deep learning

2.3. Training model

Utilizing Google Colab for training a YOLO model offers a convenient and efficient approach to harness the capabilities of cloud computing for deep learning tasks. Google Colab seamlessly integrates with Google Drive, enabling direct access to files stored in your Google Drive account from within the Colab notebook environment. To initiate the training process, begin by cloning the YOLOv7 repositories. Following this, download the pre-trained model provided by an open-source platform on GitHub. Finally, execute the command displayed below to commence the training procedure. It is important to ensure that the dataset file is readily available within your Google Drive account.

```
! python train.py --device 0 --workers 4 --batch-size 4 --epochs 100 --img 640 640 --data data/custom_data.yaml --hyp data/hyp.scratch.custom.yaml --cfg cfg/training/yolov7-custom.yaml --weight yolov7.pt --name yolov7-custom
```

3. RESULTS AND DISCUSSION

This chapter examines two cases to evaluate detection accuracy. First, it analyzes the impact of camera resolution by comparing three different resolutions: 2.1, 0.9, and 0.2 MP. Second, it evaluates the system's accuracy at varying distances between the object and the camera between 40, 80, and 120 cm.

3.1. CASE 1: the effect of camera resolution on detection accuracy

Figure 2 shows the results on the monkey detection accuracy with 3 different camera pixels which are Figures 2(a) 2.1 MP, 2(b) 0.9 MP and 2(c) 0.2 MP. The same image is used and the camera distance is set at 40 cm from the image. Based on the result as shown in Figure 2, higher camera resolution does not necessarily guarantee improved detection accuracy compared to a lower resolution camera. While it may

ISSN: 2088-8708

seem logical that more pixels would provide clearer and more detailed images for the model to analyze, other factors come into play. One key factor is the quality and diversity of the training data used to train the model. The model needs to be exposed to a wide range of scenarios and variations to generalize well and accurately detect objects. Another consideration is the computational complexity associated with higher-resolution images require more processing power and memory, which can impact the efficiency and speed of the detection process. This can become a challenge, especially when real-time or near real-time detection is required. Noise is another factor to consider. Higher-resolution images may also capture more noise or unwanted artefacts, which can interfere with the detection process. Noise reduction techniques can help mitigate this issue, but it adds an extra layer of complexity to the pipeline. Lastly, it is important to address the issue of generalization. A model trained on low-resolution images may struggle to accurately detect objects in higher-resolution images. This lack of generalization can result in reduced detection accuracy when using higher-resolution cameras to capture video, but the model is fed by a lower resolution or quality dataset for training.



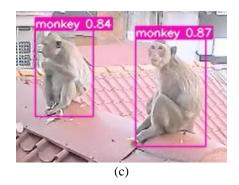


Figure 2. Monkey detection accuracy with different camera pixel quality (a) 2.1 MP, (b) 0.9 MP and (c) 0.2 MP

In conclusion, while higher camera resolution can offer benefits in certain scenarios, it is not the sole determinant of detection accuracy. The quality and diversity of training data, computational complexity, noise, and generalization issues all influence the overall accuracy of object detection systems. Optimizing the entire detection pipeline, considering these various factors is crucial for achieving accurate and reliable results.

3.2. CASE 2: the effect of camera distance on detection accuracy

The experiment's findings reveal an interesting trend: as the camera distance increases, we observe a noticeable decrease in the accuracy of object detection as shown in Table 1. This means that when the camera is placed further away from the target object, the ability of the detection algorithm to accurately identify and classify objects diminishes. One possible explanation for this phenomenon is the diminishing visibility of the target object. As the camera moves further away from the object, it appears smaller in the captured image, occupying fewer pixels. This reduction in object visibility poses a challenge for the detection algorithm, making it more difficult to precisely detect and locate the object within the frame.

Furthermore, the loss of fine details contributes to the decline in detection accuracy. With an increased camera distance, the image may lack the intricate textures, patterns, and smaller features that assist in accurate object detection. These finer details become less distinguishable, leading to misclassifications, or

even missed detections by the algorithm. Additionally, the increase in camera distance introduces potential issues such as noise, blurring, and distortion. Factors like atmospheric conditions, limitations of the camera lens, or image compression can contribute to these problems. The presence of noise and distortion hampers the detection algorithm's ability to correctly identify and classify objects, further diminishing the overall accuracy. In essence, these results highlight the critical importance of finding the optimal camera distance for achieving the highest accuracy in object detection. It becomes crucial to strike a balance where the target object is visible enough, the essential details are preserved, and noise levels are minimized. Understanding this trade-off is vital when considering the placement of cameras in real-world applications that rely on accurate object detection.

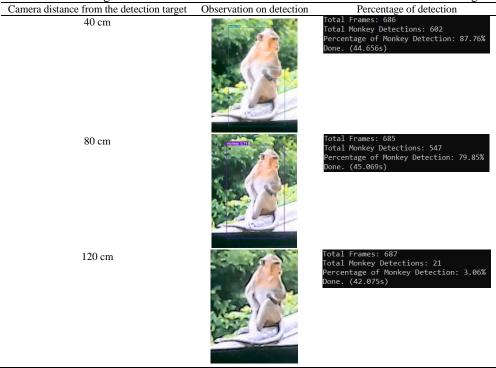


Table 1. Percentage of detection with different camera distances from the detection target

In summary, our experiment establishes a definitive correlation between camera proximity and the precision of object detection. These results underscore the importance of careful camera positioning to achieve maximal performance. By acknowledging the impact of camera distance on detection accuracy, we can make more informed decisions in practical scenarios where precise object detection is crucial.

CONCLUSION 4.

In conclusion, the trained model has demonstrated its capability to detect monkeys effectively. A crucial factor contributing to the success of deep learning approaches in monkey detection is the availability of large, annotated datasets containing diverse monkey images and videos. These datasets have played a pivotal role in improving the model's performance and generalization abilities. Despite utilizing a relatively small dataset of only 363 images, the trained model exhibited commendable performance in detecting monkeys.

However, it is important to acknowledge the limitations of working with a small dataset. Such limitations can hinder the model's ability to generalize effectively to unseen data. Consequently, this may lead to suboptimal performance on test or validation datasets, as well as in real-world scenarios. When training a machine learning model, it is essential to employ a diverse and representative dataset that encompasses the full spectrum of possible inputs and outputs the model is expected to encounter. A small dataset may lack sufficient examples of these varied combinations, thus resulting in a model that is illequipped for the task at hand. Consequently, the trained model may occasionally exhibit false positive detections and struggle to detect unseen data effectively.

FUNDING INFORMATION

This research is supported by Universiti Teknikal Malaysia Melaka and Fakulti Teknologi dan Kejuruteraan Elektrik. Also in part of Fundamental Research Grant Scheme (FRGS) under grant number, FRGS/1/2023/ICT08/UTEM/03/01 from Ministry of Higher Education Malaysia (MOHE).

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	С	Μ	So	Va	Fo	Ι	R	D	0	Ε	Vi	Su	Р	Fu	
Nur Latif Azyze Mohd	\checkmark	\checkmark	✓		\checkmark	\checkmark	✓			\checkmark	✓	\checkmark	\checkmark	✓	
Shaari Azyze															
Teow Khimi Quan	\checkmark	\checkmark	\checkmark	\checkmark		\checkmark		\checkmark	\checkmark		✓				
Ida Syafiza Md Isa			\checkmark		\checkmark		\checkmark	\checkmark		\checkmark			\checkmark		
Muhammad Afif			\checkmark				\checkmark	\checkmark	\checkmark	\checkmark					
Husman															
C : Conceptualization	I : Investigation							Vi : Visualization							
M : Methodology	R : R esources							Su : Supervision							
So : Software	D : D ata Curation								P : P roject administration						
Va: Validation	O: Writing - Original Draft								Fu : Fu nding acquisition						
Fo: Fo rmal analysis	E : Writing - Review & Editing														

CONFLICT OF INTEREST STATEMENT

All authors declare that they have no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, NLA, upon reasonable request.

REFERENCES

- M. F. Shah, "Perhilitan prefers to trap and relocate monkeys," *The Star*, 2022. https://www.thestar.com.my/metro/metro-news/2022/08/25/perhilitan-prefers-to-trap-and-relocate-monkeys (accessed May 21, 2024).
- [2] V. Janani and C. Shanthi, "Human-animal conflict analysis and management a critical survey," in *Proceedings of the 2022 11th International Conference on System Modeling and Advancement in Research Trends, SMART 2022*, Dec. 2022, pp. 1003–1007, doi: 10.1109/SMART55829.2022.10047487.
- [3] M. Vamshi, N. Marupaka, S. C. Nallamothu, and A. Naureen, "A deep learning approach monkey detection using YOLOV7," in 2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques, EASCT 2023, Oct. 2023, pp. 1–7, doi: 10.1109/EASCT59475.2023.10393326.
- [4] R. Pillai, R. Gupta, N. Sharma, and R. K. Bansal, "A deep learning approach for detection and classification of ten species of monkeys," in *International Conference on Smart Systems for Applications in Electrical Sciences, ICSSES 2023*, Jul. 2023, pp. 1–6, doi: 10.1109/ICSSES58299.2023.10199762.
- [5] I. S. B. Md Isa, C. Ja Yeong, and N. L. A. bin Mohd Shaari Azyze, "Real-time traffic sign detection and recognition using Raspberry Pi," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 331–338, Feb. 2022, doi: 10.11591/ijece.v12i1.pp331-338.
- [6] P. Malhotra and E. Garg, "Object detection techniques: a comparison," in 2020 7th International Conference on Smart Structures and Systems, ICSSS 2020, Jul. 2020, pp. 1–4, doi: 10.1109/ICSSS49621.2020.9202254.
- [7] S. Srisuk, C. Suwannapong, S. Kitisriworapan, A. Kaewsong, and S. Ongkittikul, "Performance evaluation of real-time object detection algorithms," in *iEECON 2019 - 7th International Electrical Engineering Congress, Proceedings*, Mar. 2019, pp. 1–4, doi: 10.1109/iEECON45304.2019.8938687.
- [8] R. Saraswathi, G. Shobarani, A. Subranani, and D. Tamilarasan, "Applying deep learning and transfer learning techniques for improving animal intrusion detection in agriculture farms," in 2023 International Conference on Data Science, Agents and Artificial Intelligence, ICDSAAI 2023, Dec. 2023, pp. 1–6, doi: 10.1109/ICDSAAI59313.2023.10452442.
- [9] K. Khatri, C. C. Asha, and J. M. D'Souza, "Detection of animals in thermal imagery for surveillance using GAN and object detection framework," in 2022 International Conference for Advancement in Technology, ICONAT 2022, Jan. 2022, pp. 1–6, doi: 10.1109/ICONAT53423.2022.9725883.
- [10] N. K. Bhanu and K. Sahana, "Farm vigilance: smart IoT system for farmland monitoring and animal intrusion detection using neural network," in 2021 Asian Conference on Innovation in Technology, ASIANCON 2021, Aug. 2021, pp. 1–6, doi: 10.1109/ASIANCON51346.2021.9544926.
- [11] R. Sumathi, P. Raveena, P. Rakshana, P. Nigila, and P. Mahalakshmi, "Real-time protection of farmlands from animal intrusion," in *Proceedings - 2022 IEEE World Conference on Applied Intelligence and Computing, AIC 2022*, Jun. 2022, pp. 859–863, doi: 10.1109/AIC55036.2022.9848808.

- [12] C. Ashokkumar, M. A. Kumar, R. S. Krishnan, S. M. Priya, K. L. Narayanan, and E. G. Julie, "A novel CNN-based IoT system architecture for real-time detection and prevention of animal intrusion in farmland," in *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, Sep. 2023, pp. 1355–1361, doi: 10.1109/ICOSEC58147.2023.10276300.
- [13] B. Meenakshi, S. R. Y. Aouthithiye Barathwaj, N. C. Haariharan, L. Krishnakanth, and J. Abishek, "Animal intrusion detection and ranging system using YOLOV4 and LoRa," in 3rd International Conference on Power, Energy, Control and Transmission Systems, ICPECTS 2022 - Proceedings, Dec. 2022, pp. 1–6, doi: 10.1109/ICPECTS56089.2022.10047729.
- [14] M. Shukri and M. S. Zainal, "Smart pets monitoring system using motion sensor based on IoT," Progress in Engineering Application and Technology, vol. 2, no. 2, pp. 658–669, 2021, doi: 10.30880/peat.2021.02.02.063.
- [15] E. Chandralekha, A. Muzammil Ali, V. Ritesh, and M. K. Srinivasan, "Animal intrusion detection system using SIFT features and transfer learning with MobileNetV2," in 3rd International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2023 - Proceedings, Dec. 2023, pp. 1339–1344, doi: 10.1109/ICIMIA60377.2023.10426584.
- [16] N. Mamat, M. F. Othman, and F. Yakub, "Animal intrusion detection in farming area using YOLOv5 approach," in *International Conference on Control, Automation and Systems*, 2022, vol. 2022-Novem, pp. 1–5, doi: 10.23919/ICCAS55662.2022.10003780.
- [17] E. Chandralekha, S. Ravikumar, K. Vijay, and P. Thiruselvan, "Animal intrusion detection system: protected crops and promoted safety using machine learning," in 2023 IEEE International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering, RMKMATE 2023, Nov. 2023, pp. 1–5, doi: 10.1109/RMKMATE59243.2023.10368943.
- [18] P. E. Anuvind, C. K. Abhishek, M. Shibili, C. K. Rahila, and K. Neethu, "Development and implementation of an animal intrusion detection system using image and audio processing," in 2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023, Jul. 2023, pp. 1–7, doi: 10.1109/ICCCNT56998.2023.10307574.
- [19] G. Venkatesh, G. Sai Suman, V. Sai Nikhilesh, and S. K. Ahmed, "Prevention of animal attacks on farms with IoT system," in Proceedings - 2nd International Conference on Smart Electronics and Communication, ICOSEC 2021, Oct. 2021, pp. 1–6, doi: 10.1109/ICOSEC51865.2021.9591695.
- [20] P. R. Reddy, M. V. Kumar, K. H. V. Kumari, T. Prathima, and S. Katta, "Preventing monkey menace using YOLO-based object detection model," in 2023 International Conference on Network, Multimedia and Information Technology, NMITCON 2023, Sep. 2023, pp. 1–7, doi: 10.1109/NMITCON58196.2023.10276306.
- [21] R. R. Pineda, T. Kubo, M. Shimada, and K. Ikeda, "Evaluation of the effect of transfer learning to multi-instance detection of monkeys," 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2021 - Proceedings. pp. 1357–1362, 2021.
- [22] S. M. M. Roomi, P. Rajesh, R. J. Priya, and M. Senthilarasi, "A line model based approach for monkey intrusion detection," in 2010 International Conference on Signal Processing and Communications (SPCOM), Jul. 2010, pp. 1–5, doi: 10.1109/SPCOM.2010.5560507.
- [23] W. A. Kusuma and L. Husniah, "Skeletonization using thinning method for human motion system," in 2015 International Seminar on Intelligent Technology and Its Applications, ISITIA 2015 - Proceeding, May 2015, pp. 103–106, doi: 10.1109/ISITIA.2015.7219962.
- [24] M. Harshini, M. Manaswini, N. S. Yo, K. Manideep, G. Rosy, and K. Sahithi, "Artificial intelligence based scarecrow prevention from crop destruction," in *International Conference on Sustainable Computing and Smart Systems, ICSCSS 2023 - Proceedings*, Jun. 2023, pp. 628–634, doi: 10.1109/ICSCSS57650.2023.10169360.
- [25] R. Ranjan, K. Sharrer, S. Tsukuda, and C. Good, "Effects of image data quality on a convolutional neural network trained in-tank fish detection model for recirculating aquaculture systems," *Computers and Electronics in Agriculture*, vol. 205, p. 107644, Feb. 2023, doi: 10.1016/j.compag.2023.107644.

BIOGRAPHIES OF AUTHORS





Teow Khimi Quan b S **s** obtained bachelor degree from Universiti Teknikal Malaysia Melaka in electrical mechatronic engineering in 2023. Currently, work as hardware design engineer at Exis Tech Sdn Bhd. He can be contacted at: b012010224@student.utem.edu.my.



Ida Syafiza binti Md Isa (D) SI SE received a Ph.D. degree from the University of Leeds, U.K., in 2020, worked on energy efficient access networks design for healthcare applications. She is currently a lecturer with Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. She has published several articles in this area. Her research interests include network architecture design, energy efficiency, network optimization, mixed-integer linear programming, and IoT healthcare systems. She can be contacted at email: idasyafiza@utem.edu.my.



Muhammad Afif Husman b K his B.Eng (Mechatronics) from International Islamic University Malaysia (IIUM) in 2011, MEng. (Biomedical) from University of Malaya in 2013, and Ph.D. in mechanical engineering from Leeds University UK in 2018. He is currently serving as an assistant professor in the Department of Mechatronics, Kulliyyah of Engineering IIUM. His research interests include sensor interfacing, software development and teleoperated robot communication. He is currently attached with the Center for Unmanned Technology (CUTe) IIUM, with main focus as the lead developer for IIUM Medibot, a special medical robot capable of telepresence and intelligent navigations. He can be contacted at email: afifhusman@iium.edu.my.